Challenges in evaluating technological interventions for affect regulation

Pardis Miri, Horia Margarit, Andero Uusberg, Keith Marzullo, Tali M. Ball, Daniel Yamins, Robert Flory, James J. Gross

Abstract—We evaluated whether a vibrotactile breathing pacer would influence two measures of affect during a cognitive stressor. In particular, we examined whether changes in breathing would be evident, and if so, whether these would mediate the effects of breathing pacer on self-report anxiety and skin conductance. Our results were surprising: although we observed the expected effects on breathing, we were unable to demonstrate that changes in breathing parameters were responsible for the observed changes in self-report anxiety. In addition, we failed to show that the pacer had an impact on skin conductance. In this paper, we investigate why we did not observe the expected effects. We believe our negative results have implications for evaluating technological interventions for affect regulation.

Index Terms—Haptic, Vibrotactile, Anxiety, Affect Regulation, Affect, Respiration, Slow-paced breathing, Pacer, Wearable, Linear Mixed Model, Skin conductance, Mediation, Empirical Mode Decomposition, Fast Fourier Transform, Convergent Cross Mapping

✦

1 INTRODUCTION

AFFECT regulation refers to things that we do, both cognitively and behaviorally, to alter affective states that we judge to be unwanted or context-inappropriate [1]. Examples of affect regulation include practicing slow-paced breathing, reappraisal[1], distracting oneself, smoking, or taking a nap when one is experiencing anxiety. Importantly, research has shown that some strategies are more adaptive than others (e.g., slow-paced breathing vs. smoking) [2, 3].

Repeated use of dysfunctional or unhelpful affect regulation strategies can lead to clinical conditions and diminished well being [4]. For that reason, there is considerable interest in using technology to help people more successfully regulate their affective states. In particular, vibrotactile technologies have been developed to increase our capacity to perform affect regulation in stressful situations [5–19]. For example, to reduce anxiety, PIV [17] applies slow breathing-like vibrations to the abdomen, while Doppel [12], EmotionCheck [11], and ambienBeat [18] apply slow heartbeat-like vibrations to the wrist.

Despite emerging evidence of the efficacy of affective vibrotactile devices and prototypes, the underlying mechanisms that link vibrotactile sensations to affect are not well understood. This is partly because investigating the three-step causal relationship of vibration pattern → affect regulation strategy → affect is more challenging than investigating the two-step casual relationship of vibration pattern → affect, and sometimes the link between patterns of vibrations and an affect regulation strategy is not well established or not easily quantifiable. For example, in the case of repetitive heart-beat like patterns of vibrations, it is unclear from the literature what affect regulation strategy mediated a drop in self-reported measure of anxiety, skin conductance, or heart rate variability [10–12, 15, 18, 19]. Is it response modulation, reappraisal, placebo, a combination of multiple affect regulation strategies, or something else? Answering these questions is challenging because reliance on self-reported measures of usage and success rate of known affect regulation strategies may not explain the impact of vibrations on affect. Participants may not even have noticed what affect regulation strategy they were deploying during the stressful situation.

Because of these challenges, we have focused our work around the affect regulation strategy of slow-paced breathing, which has been robustly shown to reduce skin conductance and self-reported anxiety in various contexts either with or without a stressor [20–27]. In particular, we have built a vibrotactile-based breathing pacer and demonstrated a link between the vibration patterns that guides a user through slow-paced breathing and a decrease in self-reported measure of anxiety [17].

In this paper, we extended our prior work to determine (1) whether the pacer also reduced skin conductance and (2) whether breathing modulation, induced by the pacer, would mediate the relationship between the vibrotactile patterns and the two affective measures of skin conductance and self-reported anxiety. The context for this extension was an experimental paradigm in which a group of 100 college students engaged in a set of cognitively challenging stressor tasks during the first block, and then half of them were

- P. Miri and J. J. Gross are with the Department of Psychology, Stanford University. E-mails: parism@stanford.edu, gross@stanford.edu
- H. Margarit was with the Department of Statistics, Stanford University. Email: horia@alumni.stanford.edu
- Andero Uusberg was with the Department of Psychology, Stanford University. He is currently with the department of psychology at University of Tartu. Email: andero.uusberg@ut.ee
- Keith Marzullo is a dean of College of Information Studies at University of Maryland. Email: marzullo@umd.edu
- Tali M. Ball is with Department of Psychiatry and Behavioral Sciences at Stanford University. Email: tmball@stanford.edu
- Daniel Yamins is with the Department of Psychology and Computer Science at Stanford University. Email: Yamins@stanford.edu
- Robert Flory is with Intel Labs. Email: flory@intel.com

1. Reappraisal is the cognitive reframing of an emotional event so as to have a different emotional meaning.

randomly assigned to an intervention condition with slow-paced breathing and half them were not. During this entire sequence we assessed their breathing and skin conductance and repeatedly asked about their self-reported anxiety. We expected a two-step casual relationship between the pacer and skin conductance, as well as a three-step causal relationship in which changes in breathing behavior, guided by the pacer, would explain reduced self-reported anxiety and skin conductance.

In the sections that follow, we describe the analyses we performed. We then discuss possible explanations for why the hypothesized effects were not evident. We conclude with recommendations regarding experimental design for measuring the effectiveness of affect regulation assisted by technology. To the best of our knowledge, this is the first paper that presents an in-depth examination of the mechanism that links a vibrotactile breathing pacer to affective measures of skin conductance, breathing, and self-reported anxiety. It is also the first paper that discusses the difficulties in designing experiments that measure the effectiveness of technology-based affect regulation interventions.

The specific contributions of this paper are as follows:

- We report the results of a mixed-design pre-registered experiment to evaluate the effect of a breathing pacer on psychophysiology measures.
- We interrogate the role of breathing behavior in changing affect via a set of repetitive patterns of vibrations produced by a breathing pacer.

2 HYPOTHESES

The confirmatory analyses are as follows:

1. *Efficacy of pacer in reducing breathing rate.* We hypothesized an interaction effect between group (treatment and control) and condition (Stressor 1 and Stressor 2) for the dependent variable of estimated breathing rate. We expected to observe no change from Stressor 1 to Stressor 2 in the control group, but a lower breathing rate in the treatment group during Stressor 2 when the pacer was active, which would indicate that participants were successful in following the pacer despite being cognitively engaged in performing stressful tasks.

2. *Efficacy of pacer in reducing skin conductance.* We hypothesized an interaction effect between group (treatment and control) and condition (Stressor 1 and Stressor 2) for the dependent variable of skin conductance. We expected to observe no change in the control group, but a lower skin conductance level in the treatment group, from Stressor 1 to Stressor 2. This is because we expected that slow-paced breathing deployed by the treatment group in the Stressor 2 condition when the pacer was active would be more effective in reducing skin conductance compared to other forms of affect regulation they may have deployed in Stressor 1.

Conditional upon observing the effect of the pacer on reducing breathing rate (confirmation of hypothesis 1), reducing skin conductance (confirmation of hypothesis 2), and previously observed reduction in self-reported anxiety, we had the following confirmatory mediational hypotheses.

3. *Indirect effect of pacer in reducing skin conductance through breathing rate.* We hypothesized that the breathing rate should mediate the impact of the pacer on changes in skin conductance measure from Stressor 1 to Stressor 2.

4. *Indirect effect of pacer in reducing anxiety through breathing rate.* We hypothesized that breathing rate should mediate the impact of the pacer on the change (in this case drop) in self-reported measure of anxiety (STAI-6) from Stressor 1 to Stressor 2.

The exploratory hypothesis was envisioned to test whether all participants in the treatment group equality benefited from the breathing manipulation. If not, we were interested whether the effects of interests (hypotheses 1–4) were evident in a subset of participants who were most effective in their breathing manipulation.

3 RELATED WORK

3.1 Short- and long- term benefits of slow-paced breathing

Slow-paced breathing within the range of 4 to 8.5 breaths per minute has been applied as a non-pharmacological method in stress management and biofeedback training. In general, during slow-paced breathing, inhalation temporarily suppresses parasympathetic activity and causes an immediate increase in heart rate, while exhalation decreases the heart rate and restores parasympathetic activity. This results in greater heart rate variability than the resting state. When an individual breathes at their resonance frequency [20] (which is unique for each individual and typically falls in the range from 4.5 to 7 breaths per minute), this effect is magnified significantly. Additionally, prolonged exhalations at this rate reduces physiological arousal quickly (within 60 seconds). In addition to immediate benefits, there have been long term benefits on cardiovascular and respiratory functions. For example, practicing slow-paced breathing for 20 minutes every day for a few months has been shown to decrease blood pressure and help with hyperventilation. In addition, as we age, our lung capacity and functioning decrease, and practicing slow deep breathing regularly can slow these gradual decreases. [20–26]

Although it is recommended for an individual to practice slow-paced breathing at their particular resonance frequency for optimal benefits, both Leher et. al and Khazan's protocols [28, 29] require multiple sessions over a course of multiple weeks which render them impractical for an in-lab study. For that reason, in PIV [30], we introduced an approximation protocol, individualized to each participant, to find a breathing rate close to their resonance frequency that reduced skin conductance and increased heart rate variability. In this study, we use this protocol to tune our pacer to guide the treatment group through slow-paced breathing.

3.2 Vibrotactile technologies for affect regulation

Compared to other studies evaluating vibrotactile technologies for affect regulation, this study measured a greater range of psychophysiological indicators, which allowed us to assess more comprehensively the mechanism of the breathing pacer on affect change. Besides our pacer, several other devices have been studied that use vibrotac-
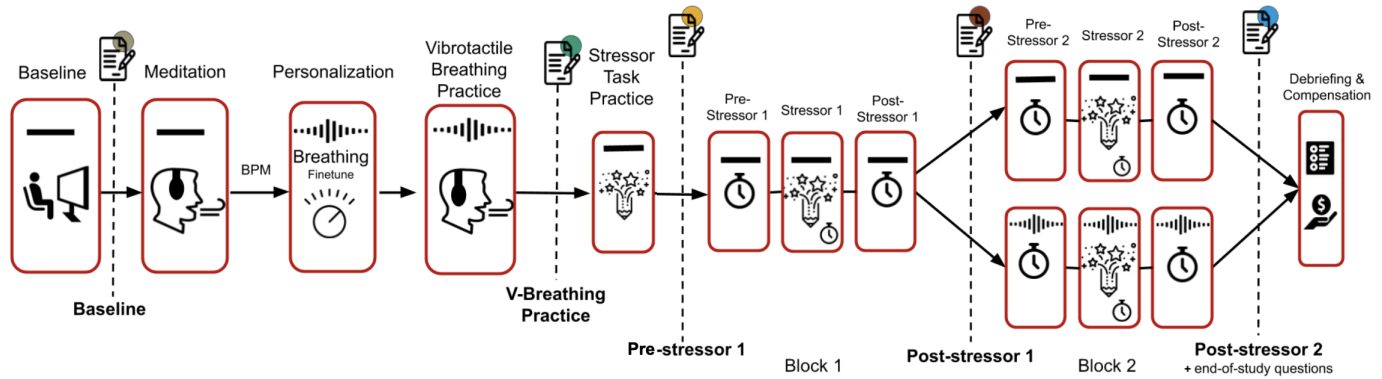
Fig. 1: Procedure flowchart of the study. All participants went through the same procedure until end of Post-stressor 1. At that point, participants who were randomly assigned to the control group went through Block 2 while receiving no vibrations (indicated by a solid black line), while participants who were randomly assigned to the treatment group go through Block 2 while receiving vibrations. Subjective measures of anxiety (i.e., STAI-6) were collected at five conditions of Baseline, V-Breathing Practice, Pre-stressor 1, Post-stressor 1, and Post-stressor 2 (indicated by a color-coded pen-and-paper icon). Physiological (i.e., skin conductance and respiration) measures were collected during the entire procedure. [17]

tile technology for affect regulation. These include[2]: Doppel [12], BoostMeUp [15], and ambienBeat [18], all three of which provide heartbeat-like stimulation on the wrist; Spire Stone [14], which monitors users' respiration and delivers interventions through an app; and Just Breathe [10], Breath Booster [31], Calm Commute [32] a series of in-car haptic and audio interventions for use while driving. Of the studies used to evaluate the other vibrotactile technologies mentioned, only two analyzed skin conductance as a measure of anxiety (Doppel, Just Breathe), and one (Just Breathe) failed to show a reliable change in skin conductance. In addition, only two analyzed respiratory data in any form (Just Breathe and Calm Commute) and one (Calm Commute) used measures of Heart Rate Variability to justify a calming effect [3]. Furthermore, none of these studies analyzed breathing modulation in terms of breathing regularity, percentage of breathing modulation, or chest-to-abdomen displacement ratio, all of which are known to have a link to change in affect [29, 33]. Lastly, none of these studies explored a three-step causal relationship (vibration pattern → affect regulation strategy → affect). Our study used the measured skin conductance and breathing data, and also derived measures of breathing behavior (e.g., breathing rate, irregularity, and the extent to which treatment group followed the pacer). We were able to analyze the effect of using the breathing pacer on each measure and also develop and analyze an explanatory hypothesis about the affect regulation mechanism of the breathing pacer by testing for breathing mediation effect to explain the pacer's calming effect.

2. For a comprehensive review of vibrotactile devices for affect regulation, see Background sections in [17, 18]. In this paper we are only focused on those devices whose effect on psychophysical measures have been reported.

3. Heart Rate Variability (HRV) and breathing rate are markers for two different organ systems, namely the cardiovascular system and the respiratory system. But, the markers are linked: short-term increase in HRV occurs while breathing near the resonant frequency. Hence, we believe that HRV should be used as a manipulation check on slow-paced breathing rather than as a measure of affective outcome.

4 METHODS

In this section, we briefly describe the study design. An extensive explanation can be found in [17].

4.1 Participants

Participants were recruited via a university pool of Stanford undergraduate students with the incentive of either course credit or hourly payment for their time. Out of the total of 100 volunteers, 3 had to be excluded due to procedural errors. 44 participants (29 female, 15 male) were randomly assigned to the treatment group and 53 (32 female, 21 male) to the control group. The IRB review and approval prior to and consents from all the participants during the study were obtained.

4.2 Protocol

Before starting the study, participants, with the help of research assistants, placed a few devices on different locations of their body. Based on the recommendations from [30], each participant had two tactors taped to their abdomen. They were also fitted with breathing gauges around the chest and abdomen and electrodermal activity sensors placed on the index and ring fingers of their non-dominant hand. Participants were given noise-cancelling headphones to aid in blocking out any potential vibration sounds.[4]

The study consisted of the following stages: Baseline, Meditation, Personalization, Vibrotactile Breathing Practice, Stressor Task Practice, Block 1, and Block 2 (see Figure 1). All participants followed the same procedure until the end of Block 1, at which point the treatment group received the breathing pacer intervention while the control group did not. Below, we summarize the stages of the protocol and

4. C2 tactors used in this study can make audible noise in a quiet lab environment when vibrating at frequencies over 165Hz [34] without noise canceling headphones. In the personalization procedure, C2 tactors were not driven beyond 180Hz. Beyond this frequency, C2 tactors vibrations become audible even when noise canceling headphones are in use.
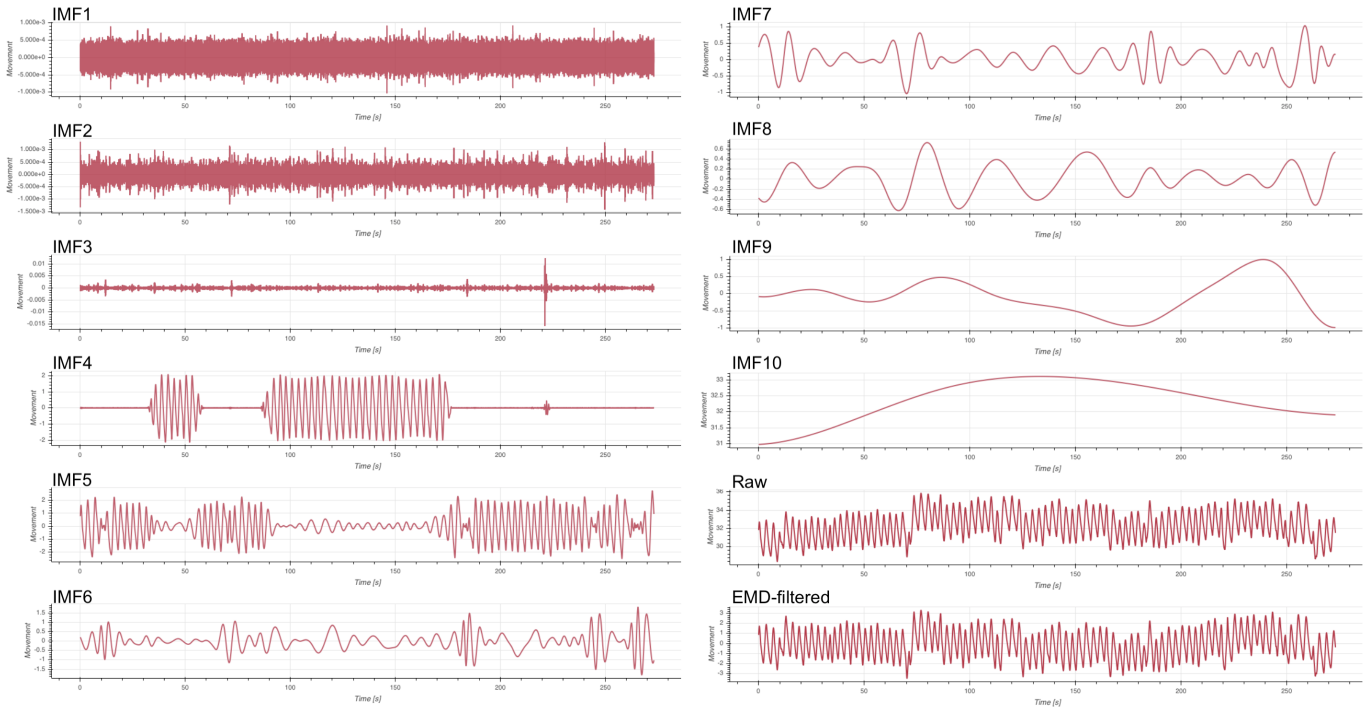
Fig. 2: An example of empirical mode decomposition (EMD) approach to decompose a respiratory signal into nine intrinsic mode functions (IMFs) components. We removed the first two and the last IMF to filter tissue artifacts (e.g., any muscle movement), posture change, and background noise from the respiratory signal.
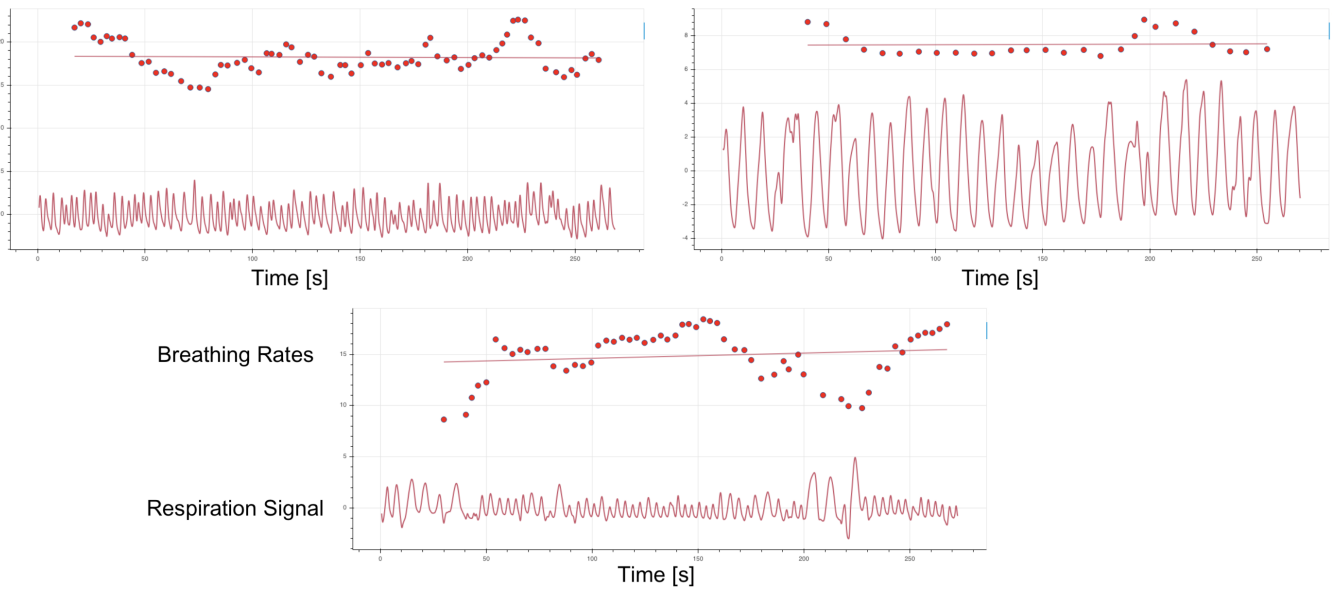

Fig. 3: Examples of breathing signals. The red dots are rolling average breathing rates with window of five breaths, and the red line shows the linear regression for the breathing rates. **Top left:** Passive breathing. **Top right:** Active breathing. **Bottom:** Combination of both passive and active breathing. All three examples were from participants in the treatment group who received pacer intervention during the Stressor 2 condition. The cutoff to categorize a breathing rate as active was 14 breaths per minute or slower.

describe the stressor task used in Blocks 1 and 2. For more detailed information about the protocol and the stressor task, see the Methods section of Miri et al. [17].

In the Baseline stage, participants watched a 5-minute video of a pipe changing colors and were asked to count how many times the pipe turned green; the objective of this stage was to obtain baseline physiology measurements. In the Meditation stage, participants listened to a 5-minute audio guiding them through slow-paced abdominal breathing techniques. The objective of this stage was twofold: to train
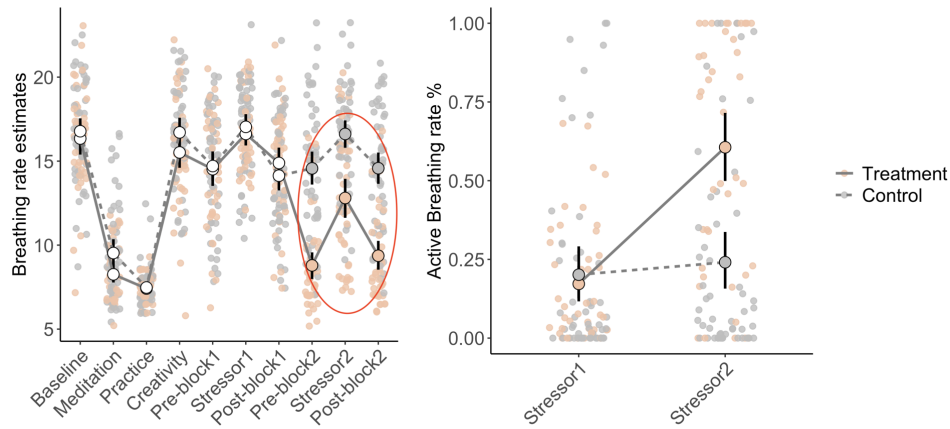
Fig. 4: **Left:** Average breathing rate estimate during conditions of the study for both treatment and control groups. An interaction effect was observed between group (treatment and control) and condition (Stressor 1 and Stressor 2): the treatment group receiving vibrotactile patterns during Stressor 2 experienced a drop in the breathing rate compared to the control group. The solid line indicates the treatment group; the dotted line indicates the control group. Note that the treatment group breathing was significantly lower compared to the control group when the pacer was active (highlighted by the red circle). **Right:** Average active breathing rate during Stressor 1 and 2 conditions as reported on a scale of 1 to 100 by both treatment and control groups. An interaction effect was observed between group (treatment and control) and condition (Stressor 1 and Stressor 2): we did find that receiving vibrotactile patterns had an influence on portion of active breathing.
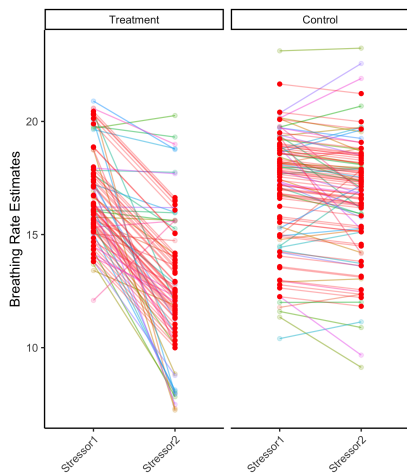

Fig. 5: Linear mixed model fit to breathing rate estimates. Note that the model fit lines with low negative slopes for the control group and lines with large negative slopes for the treatment group (shown in red), in both cases considering individualized intercepts. The model predicts a significant drop in breathing rate for the treatment group going from the Stressor 1 to Stressor 2 condition.

participants to prepare them for synchronous breathing with the vibrations, and to estimate their slow-paced respiratory rate in breaths per minute. This is the personalized respiratory rate that was used to guide the users through vibrotacticle slow-paced breathing for the rest of the study.

In the Personalization stage, we established personalized frequencies and amplitude of vibrations for each participant using a staircase method. We determined two frequencies of vibrations to allow participants to distinguish between the two phases of inhalation and exhalation. Typically, these two frequencies were approximately 30 to 40 Hz apart to allow them to be easily distinguished from each other. Research assistants worked with participants to arrive at these personalized frequencies and amplitudes. The details of the personalization protocol can be found in [30].

In the Vibrotactile Breathing (V-Breathing) Practice stage, participants practiced synchronizing their breathing with the personalized vibrations for 90 seconds. This stage allowed the research assistants to determine whether the choices for pace, amplitude, and frequencies allowed the participant to breathe slowly while effortlessly in sync with the vibrations. If not, then the research assistant would adjust these values so that breathing was effortless.

In the Stressor Task Practice stage, participants were introduced to the compound remote associate (CRA) task to be used as a stressor. Compound remote associate questions consist of three English words whose answer is a fourth word that is associated with all of them. For example, the answer to the CRA question "river/note/account" would be "bank." The specific CRA questions were taken from Bowden & Jung-Beeman [35]. Block 1 consisted of the first stressor and two waiting periods occurring before and after the stressor. Waiting periods, called Pre-stressor 1 and Post-stressor 1, allowed two minutes to pass between stressors so that any anxiety from the stressor would not affect measurements taken at other stages. The stressor task was 4 minutes long and consisted of 27 CRA questions with a time limit of 9 seconds each. Running out of time on a question was equivalent to an incorrect answer. If a question was answered correctly, participants would still have to wait the full 9 seconds before seeing the next question; this was done to ensure that all participants spent an equal amount of time in the stressor task. During Block 2, which was otherwise
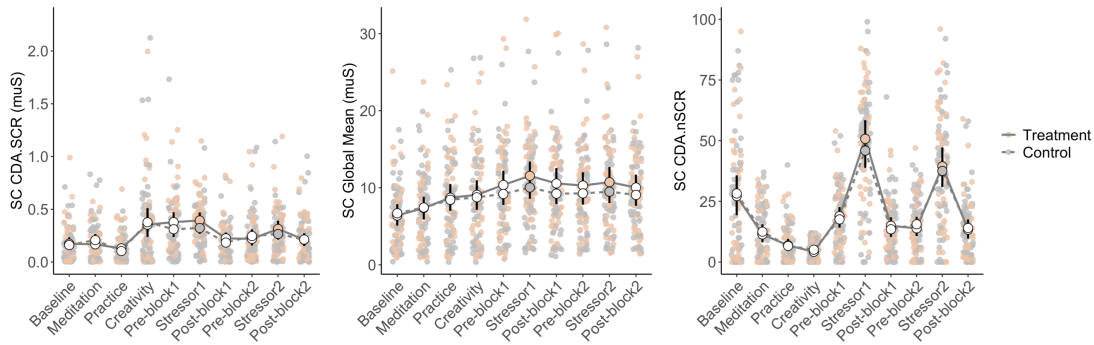
Fig. 6: Three measures of skin conductance all indicating no difference between the treatment and control group. **Right:** The large inferred number of peaks in the phasic drive (CDA.nSCR) during Stressor 1 and Stressor 2 demonstrated that the participants experienced significantly higher levels of sympathetic arousal during the stressor tasks
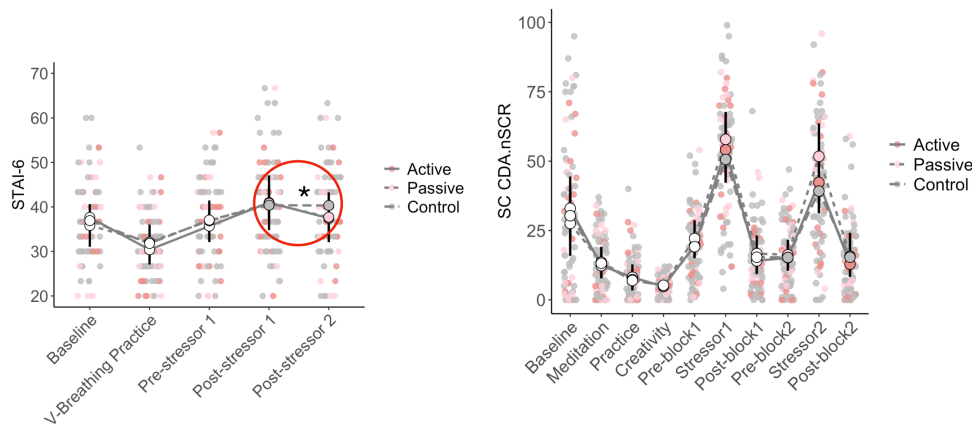

Fig. 7: **Left:** Average anxiety self-reported (STAI-6) rate estimate during conditions of the study for both treatment and control groups. Two two-way interaction effects was observed between group (passive, active, and control) and condition (Stressor 1 and Stressor 2): both passive and active groups receiving vibrotactile patterns during Stressor 2 experienced a drop in their self-reported anxiety compared to the control group. The solid line indicates the active group; the dotted lines indicate the passive (in pink) and control (in gray) groups. Note that both passive and active groups anxiety were significantly lower compared to the control group when the pacer was active (highlighted by the red circle) which may suggest an existence of a placebo effect. **Right:** Average skin conductance (specifically CDA.nSCR) during Stressor 1 and Stressor 2 conditions as reported on a scale of 1 to 100 by passive, active, and control groups. No interaction effect was not observed between group (passive, active, and control) and condition (Stressor 1 and Stressor 2): Passive, active, and control groups were indistinguishable.

identical to Block 1, the treatment group received ongoing personalized vibrations and were instructed to conform their breathing with the sensation of the pacer.

5 MEASURES

We used Qualtrics to collect self-reported measures and FlexComp System with BioGraph Infiniti Software T7555M by Thought Technology to acquire all physiology measures in this study.

5.1 Respiration measures

Our primary respiratory measure was breathing rate (number of breaths per minute). In addition to this primary outcome, to determine compliance with the pacer, we calculated active breathing percentage which reflects the extent to which the treatment group followed the pacer in a response

window. For exploratory purposes, we also computed additional breathing indices including Breathing irregularity (measured through coefficient of variance) and chest to abdomen displacement ratio (slope of a regression line fitted over a Konno-Mead plot [36, 37]).

5.2 Skin Conductance

Skin conductance (SC) is a measure of the degree to which the skin transmits a small electric current between two electrodes, and it reflects sympathetic arousal (SA: i.e., activation of the sympathetic nervous system), which is associated with the so-called fight-or-flight response, as the body anticipates and prepares for action, either mental or physical.

5.3 Self-reported anxiety

To measure anxiety, we collected responses to the State-Trait Anxiety Inventory (STAI-6) after Baseline, V-Breathing

Practice, Pre-stressor 1, Post-stressor 1, and Post-stressor 2 (indicated by a color-coded pen-and-paper icon in Figure 1). We chose STAI-6 – the brief six-item version of the State-Trait Anxiety Inventory – over the more commonly used STAI-20 because it reduced the burden on the participant. In addition, STAI-6 scores correlate highly with scores on the full-length STAI-20 ($r = 0.95$) [38]. The results of this measure were previously reported in [17].

5.4 Self-reported Regulation

At the end of the study, we asked open-ended questions about the type of affect regulation participants used during block 1 and block 2.

6 DATA PROCESSING

6.1 Respiratory data

Before extracting respiratory measurements of interest (mentioned above), we used an empirical mode decomposition (EMD) based signal processing technique, inspired by Liu et al. [39], to remove tissue artifacts (e.g., any muscle movement), posture change, and background noise from respiratory signals. In addition, we applied a smoothing technique using a rolling window of $\frac{2}{3} \times$ sampling frequency (in our case 256 Hz). Below, we describe in more detail our approach to artifact removal.

6.1.1 Artifact removal of respiration signal

Empirical Mode Decomposition (EMD) [40] decomposes a time-varying signal into a sum of finite components, referred to as intrinsic mode functions (IMFs), each of which represents an oscillatory mode (see Figure 2). A critical step in achieving an artifact-free signal is selecting the appropriate IMFs that contribute to the signal and excluding those IMFs that mainly contribute to the contamination of the signal. In the context of a respiration signal derived from ribcage or abdomen displacement, tissue artifacts (e.g., any muscle movement unrelated to breathing), and background noise are usually present in higher frequency bands (i.e., several first IMFs) while lower frequency bands (i.e., several last IMFs) often correspond to artifacts resulting from posture changes. By removing these high and low frequency bands corresponding to such artifacts, an artifact-free respiratory signal can be reconstructed by summing up the remaining IMFs.

We investigated different applications of EMD for removing artifacts ([39, 41, 42]), each applied to differing ways the breathing signal was captured. We found none of them worked well with our signal. Instead, we found that removing the first two IMFs (which always corresponded to tissue artifacts) and the last IMF (which always corresponded to the participant's posture change). This resulted in less artifact removal, but enough to estimate breathing rate using both peak finding and Fast Fourier Transform (FFT) approaches. We found that the removal of the lowest computed IMF was more important than the first two highest computed IMFs in measuring breathing rates, as described next.

6.1.2 Breathing rate

For the purposes of confirmatory analyses, we chose a simple and robust approach for measuring breathing rates. After removing artifacs using the generic EMD approach described above, we identified all peaks and troughs [5] We then computed a set of breathing rate averages using a sliding window of 5 breaths, and returned the median of this set as the dominant breathing frequency. The window size of 5 was chosen because we did not find a significant difference between what this algorithm computes versus what an FFT-based approach, described in Appendix A, computed for segments over which it returned a valid result. This appendix section also explains why we did not use the FFT-based approach.

6.1.3 Degree to which the treatment group followed the pacer

For the purposes of exploratory analyses which were designed to assess whether participants in the treatment group varied in benefits they enjoyed from the pacer, we computed a measure of how well that participant followed the pacer. We computed the set of breathing rate averages using a sliding window of 5 breaths, and partitioned this set into two, using the cutoff value of 14 breaths per minute to distinguish passive from active breathing rates [29]. Then, we computed the ratio of active breathing rate averages to the total number of averages in the original set. Thus, those who followed the pacer fully — and were thus actively breathing — were assigned a measure of 1, those who ignored the pacer were assigned 0, and those remaining were assigned a value between 0 and 1. Then, based on visual inspection of Figure 4, we further classified a participant with an active breathing ratio > 0.85 as an active participant, less than 0.4 as passive, and the remaining as in-between participants. This resulted in 17 participants in active, 16 in passive, and 11 in in-between groups. The in-between group members could not fully shift their breathing to match the pacer rate at all times and so their attempts led to a greater breathing inconsistency compared to the passive and active groups. In contrast to the slow, deep, and diaphragmal breaths associated with proper slow paced breathing chemistry, inconsistent breathing can be linked to negative affective states such as anxiety [29]. Because of this, in some of our exploratory analysis, we excluded the in-between group participants to focus on testing the distinction between passive, active, and control groups who all exhibit approximately consistent breathing patterns. Figure 3 shows examples of participants belonging to each of these categories. Note that in addition to the difference in breathing rate, active breathing also resulted in a larger breathing wave amplitude as compared to passive breathing.

We believe that whether participants followed the pacer or not was not as important as whether they managed to slow down their breathing sufficiently enough to have a beneficial effect. When we did directly compare the pacer phase with the breathing wave, we observed misalignments

5. To extract the peak-related information, we used Python `find_peaks` with the following settings: `find_peaks(signal, distance = 2 × sampling_frequency, width = .8 × sampling_frequency, prominence = .3)`.

during the stressor phases. For example, we would frequently see a participant take a deep slow breath when the pacer was more than half way through the inhalation phase. We think it was important to count this slow breath as an attempt to follow the pacer and so a model of how much participants aligned their breathing with the pacer was not contextually appropriate to use.

6.1.4 Efficacy of the pacer in reducing breathing irregularity

For the purposes of exploratory analysis we measured the irregularity in breathing depth, which is the coefficient of variance (CV) of the respiration signals. This measure is known as the relevant standard deviation expressed as a percentage of a signal mean (i.e., $\frac{SD}{Mean} \times 100$). A higher CV corresponds to more irregularity in the depth of breathing.

6.1.5 Efficacy of the pacer in reducing chest to abdomen displacement ratio

For the purposes of exploratory analysis, we examined ribcage-abdominal motion using the method of Konno and Mead [36, 37], where the extensions of the abdomen and rib cage are displayed on an XY plot. There is no axis corresponding to time on a Konno-Mead plot; instead, the breathing waves together generate a trace on the Cartesian plane. When the chest and abdomen move relatively in synchronization, the trace is a figure that slopes upwards and to the right. We then fit in an ordinary least squares (OLS) regression line which is a measure of the ratio of abdomen to chest motion: slopes less than 1 result from more abdominal compared to chest breathing. Appendix C explores this approach in more detail.

6.2 Skin Conductance Data

For the purposes of confirmatory analyses, we used two general approaches for skin conductance analysis: a model-based approach (called CDA, for continuous decomposition analysis) which uses mathematical models to map between observed SC and SA, and a model-free approach, such as the average of an observed SC within a certain time frame. Ledalab [43] is a widely used package that infers SA from observed SC. We used two model-based measures produced by Ledalab: `CDA.nSCR`, which is the total number of significant peaks in the phasic (fast-changing) drive, and `CDA.SCR`, which is the average phasic driver value in a response window.[6] We also used the model-free measure `SC.mean`, which is the averaged observed skin conductance over a response window. See the Appendix B for details on how the LedaLab package computes `CDA.nSCR` and `CDA.SCR`.

7 Data Analyses

To test the confirmatory analyses of (1) efficacy of pacer in reducing breathing rate and (2) efficacy of pacer in reducing skin conductance, we ran linear mixed models to test the

---

6. For replicability purposes, here is a sample invocation of the Ledalab script with a skin conductance sampling rate of 256 Hz: `Ledalab('path_to_data', 'open', 'text', 'downsample', +2, 'analyze', 'CDA', 'optimize', 2, 'export_era', [6 240 .05 1])`. The response window is 6 to 240 seconds, and the threshold for SCR peaks is .05 microsiemens.

interaction effect between group (treatment and control) and condition (Stressor 1 and Stressor 2) on the dependent measures of interest while controlling for stable individual differences. Models were implemented in R with the `lme4` package according to the following specifications: `lmer(DV ~ condition * group + (1 | id))`. The effect confidence intervals were bootstrapped.

To test the confirmatory analyses of (3) indirect effect of pacer in reducing skin conductance through breathing rate and (4) indirect effect of pacer in reducing anxiety through breathing rate, we used the `mediate` function of the `mediation` package in R (see AppendixG for sample code). To assess the impact of compliance with our instructions in the treatment group, we quantified the degree to which the treatment group adhered to the pacer and classified the treatment participants into three groups of active (largely followed the pacer), passive (largely ignored the pacer), and in between (which occasionally followed the pacer: we excluded this group from further analysis). We then tested for the interaction effect between group (passive, active, and control) and condition (Stressor 1 and Stressor 2) for the two affect measures of skin conductance and self-reported anxiety.

In addition to these four confirmatory analyses and one exploratory analysis, we also conducted some additional analyses to better understand the patterns of findings that emerged from these analyses.

8 Results

During the personalization procedure, we tailored the vibrotactile breathing patterns to the needs of each individual participant. The slow breathing pace was in the range of 5.5 to 9 with median of 7 and mean of 7.15. Below we discuss the details of our findings.

8.1 Efficacy of the pacer in reducing breathing Rate

In the first confirmatory analysis, we ran linear mixed models on the measure of breathing rate and found an interaction effect between the condition (Stressor 1 and Stressor 2) and group (treatment and control). Consistent with our hypothesis, the treatment group's breathing rate dropped from Stressor 1 to Stressor 2 as a result of the pacer intervention (see Figure 4, left, and Figure 5). The observable large effect (intercept = 17.02, $\beta = 3.27$, $p < .0001$, CI = $[2.18, 4.36]$) suggests that on average, participants lowered their breathing rate.

Some participants were more successful than others in implementing slow-paced breathing throughout the Stressor 2 window. The subsection below describes follow up exploratory analyses on this in more details.

8.1.1 Exploratory analyses: Efficacy of the pacer in increasing ratio of active breathing

To quantify how successful treatment group participants were in following the pacer, we derived a measure of active breathing percentage, as explained in Section 6.1.3. Figure 4, right, shows the ratio of active breathing in two conditions of Stressor 1 and Stressor 2. During Stressor 1, the participants (treatment and control groups combined) performed

an average of 19% of active breathing.[7] During Stressor 2, the proportion of active breathing percentage increased to approximately 57% for the treatment group when the pacer was active, but did not significantly change for the control group.

While the average breathing rate in the treatment group dropped significantly during Stressor 2, the majority of the participants were not able to breathe consistently at a rate slow enough to be classified as following the pacer. Instead, we observed a significant variability in the breathing rates, suggesting that participants were more likely to switch back and forth between passive breathing when solving a problem and active breathing when they were able to attend to the pacer. This pattern of breathing resulted in more irregularity in breathing data, and every time a switch between passive to active breathing occurred, more chest breathing became involved.[8]

8.2 Exploratory analyses: Were participants' performance compromised due to how much they engaged with the pacer?

Stressors and stressful situations come in various types, ranging from low to high cognitive demand, as well as anticipatory, in-the-moment, or ruminating stress. In the case of the CRA task, it is highly cognitively demanding, in-the-moment stressor.

We chose the CRA task because of its ecological validity. People often become anxious while engaging in cognitively demanding activities. In addition, the CRA task was well suited to our mixed design experiment. Our mixed design required two stressor blocks, and so the stressor must be repeatable.Lastly, the CRA tasks don't have a learning effect: one doesn't get better at answering the CRA-generated questions.

Our findings suggest that the pacer was usable with the CRA stressor. In particular, the degree of following the pacer did not affect CRA performance. For instance, the CRA scores of the passive group did not differ significantly from the average score (intercept = 10.13, β = -0.73, p = .50, CI = $[-2.85, -1.48]$). Likewise, the active group didn't have a score any higher or lower than the average score (intercept = 10.47 , β = -1.18, p = .24, CI = $[-3.17, .80]$). Therefore, following or ignoring the pacer did not result in boosted or reduced CRA task performance.

8.3 Efficacy of the pacer in reducing skin conductance

In the second confirmatory analysis, we ran linear mixed models on three measures of skin conductance: CDA.SCR, CDA.nSCR, and CS.mean. Although we only pre-registered our hypothesis for the measure CDA.SCR, when we did not observe the expected interaction effect, we used the other two measures to ensure that none of the three measures

7. That is, of the breaths taken in this condition, 19% were at a rate slow enough to be considered to be active breathing. Because of the rolling average used to compute the breathing rates, the first 4 breaths were not included in this percentage.

8. See Appendix C for analysis of pacer effect in modulating breathing behavior in terms of breathing irregularity (coefficient of variance), and ratio of abdomen to chest displacement (slope of a regression line fitted over Konno-Mead plot).

showed the hypothesized effect. We did not find an observable interaction effect between group (treatment and control) and condition (Stressor 1 and Stressor 2) for CDA.SCR (intercept = .39, β = .02, p = .36, CI = $[.02, -.001]$), CDA.nSCR (intercept = 54.86, β = 4.45, p = .14, CI = $[4.45, .11]$), or CS.mean (intercept = 11.55, β = .28, p = .24, CI = $[-.21, .70]$) (see Figure 6, center and right). That is, we did not find a significant drop in skin conductance (reflecting a drop in sympathetic arousal) as a result of the breathing pacer intervention using either model-based or model-free measures of skin conductance. The treatment and the control groups are indistinguishable in terms of skin conductance measures in Figure 6).

Contrary to our hypothesis, the results suggest that although the treatment group participants attempted to use slow-paced breathing, doing so was not any more or less effective than whatever affect regulation method they may have deployed during the Stressor 1 condition during which there were no pacer vibrations (within-subject comparison) or whatever the control group participants deployed (between-subject comparison).

8.4 Indirect effect of pacer in reducing skin conductance through breathing rate

In the third confirmatory analysis, because the two-step causal link between the pacer and skin conductance was not observed (see subsection above), we were unable to show that there exists a three-step link between these two measures mediated through breathing rate.

8.5 Indirect effect of pacer in reducing anxiety through breathing rate

In the fourth confirmatory analysis, we failed to find the mediation effect of the breathing rate to explain the link between the pacer and self-reported anxiety. We previously found an interaction effect between group (treatment and control) and condition (Stressor 1 and Stressor 2) for STAI-6, which suggested a calming effect of the pacer in reducing stress (see [17], Results and Discussion section). We tested whether this effect can be explained by the breathing rate, and did not find the average causal mediation effects (ACME) = .03, CI = $[-1.69, 1.71]$, p = .97; ADE = -2.45, CI = $[-5.56, .5]$, p = .09; total effect = -2.41, CI = $[-4.83, -.10]$, p = .04).

8.5.1 Did other measures of breathing modulation cause a drop in self-reported anxiety?

We have shown that the pacer modulates breathing behavior and that the breathing behavior was causally linked to self-reported anxiety. But we have failed to show that the pacer produced a reliable change in self-reported anxiety through breathing modulation. This has led us to further test whether the other measures of breathing we collected (see Appendix C, i.e., breathing regularity, chest to abdomen breathing ratio, and active breathing ratio) could explain the pacer's effect as a mediator. Unfortunately, none of them showed a mediation effect to explain the drop in self-reported anxiety (STAI-6). This suggests that identifying more measures of breathing behavior may not be a fruitful direction to pursue. Instead, we may benefit more by

focusing on improving methods for collecting self-reported anxiety measures.

8.6 Efficacy of the pacer in reducing anxiety and skin conductance in subset of treatment group participants

In the exploratory analysis, after observing that not all participants in the treatment group equality benefited from the breathing manipulation (see section 8.1.1), we tested whether the effects of the pacer was evident in a subset of participants who were most effective in their breathing manipulation. Figure 7 (Left) shows that participants in both active and passive groups both reported benefits from the pacer as compared to the control group. However, the control, active, and passive groups were indistinguishable in terms of skin conductance during Stressor 2 Figure 7 (Right).

9 DISCUSSION

9.1 Summary of Results

In this study, we found that the pacer had the intended impact on breathing but contrary to expectations, there were no effects of the paced breathing on skin conductance and breathing modulation did not mediate the effect of the pacer on either skin conductance and self-reported anxiety (reported separately). In addition, we failed to show that the effects of the pacer was evident in a subset of participants who were most effective in their breathing manipulation. In the following sections, we will consider why we might have failed to observe the expected effects and attempt to draw some conclusions and provide suggestions about future works in this area.

9.2 Why did we not observe the effect of pacer on skin conductance?

In an attempt to understand why the control, active, and passive groups were indistinguishable in terms of skin conductance, we first checked that the skin conductance measurement and the experimental design were meaningful (see Appendix E for more details). We then investigated several explanations based on the data we collected. Below, we describe two explanations that we believe are the most and one that we think is the least plausible. In Appendix F we listed other explanations that could be possible but are less data-driven as compared to the explanations presented below.

1. *The control group was not passively experiencing stress: They were spontaneously regulating.*

One possible explanation of our failure to find reductions in skin conductance during Stressor 2 is that the affect regulation spontaneously deployed by the control group was as effective as the slow-paced breathing deployed by the treatment group, thus concealing the hypothesized effect of slow-paced breathing on skin conductance.

The possibility that the control group was spontaneously engaging in affect regulation was supported by their self-report of the affect regulation they deployed during Pre-stressor 2 (a stress anticipation window prior to stressor tasks) and Stressor 2 (when they were performing the stressor task). We grouped the control participants into those

who didn't report using an affect regulation strategy in Pre-Stressor 2 and Stressor 2, and those who did use some affect regulation strategy. For some participants, it was unclear what affect regulation strategy they deployed. For example, participants s076 reported "thinking of words that I know". We included such participants in the group that used some affect regulation strategy. We identified five participants (10%) in Pre–stressor 2 and 9 in Stressor 2 (18%) who reported using no affect regulation strategy. Therefore, most the control group did not report passively experiencing stress.

2. *The active group followed the pacer at a cost: Cognitive load induced by synchronizing breathing with the pacer*

The second possible explanation as to why the treatment group may not have shown the reduction to skin conductance is because skin conductance is sensitive not only to affective responses but also to cognitive load. The reason this is relevant is that the participants in the treatment group had two tasks, whereas the control participants had only one. More specifically, the treatment group both had to perform the CRA tasks and had to follow to the pacer, while the control group only had to perform the CRA tasts. On this account, even though some of the participants showed reduced anxiety and reduced breathing rate, they may have shown equivalent skin conductance responses to the control group because of the greater effort that was expended trying to conform to the demands of the CRA tasks as well as to the pacer. We grouped the treatment participants into two: those who mentioned something about needed to context switch between breathing and stressor tasks and those who did not. Fourteen participants (33%) were in the first group, illustrating the difficulty of following the pacer while engaged in the CRA task.

3. *Not observing the impact of slow-paced breathing on skin conductance does not imply lack of its existence* We believe that not observing a difference between active and passive groups in terms of skin conductance does not contradict the well-known fact that slow-paced breathing has an effect on skin conductance. We have found evidence that the pacer modulated the average breathing rate, and we observed that breathing behavior was unidirectionally linked to temporal changes in skin conductance. Specifically, for each participant, regardless of group membership (Treatment or Control) or condition (pre-stressors or stressors), we observed a unidirectional relationship between the time series of their breathing signal (chest or abdominal displacement) and their correspondent skin conductance. We made this observation using Convergent Cross Mapping (CCM) [44], which is a method that distinguishes causality from correlation in nonseparable weakly connected dynamic systems. We tested this causal relationship and have shown that, indeed, breathing caused a change in skin conductance, and not the other way around. See Appendix D for more details on CCM analysis.

9.3 Why did we not observe the indirect effect of pacer in reducing anxiety through breathing rate?

One plausible explanation for not observing the mediation effect is because the self-reported anxiety measures were not collected immediately after the stressors sections. Instead,

they were collected at the end of the 2-minute recovery time (Post-stressor 1 and Post-stressor 2). Observe in Figure 6, right, that during the recovery time, the level of sympathetic arousal had significantly dropped for both the treatment and control groups, but there was no significant difference between the two in the level of this drop. Therefore, it is possible that the retrospective measure of STAI didn't capture the nature of anxiety that could have been captured had the questionnaires been administered immediately after the stressor conditions. Perhaps that is why we didn't observe a larger interaction effect and we failed to observe the mediation effect through breathing rate.

Another plausible explanation is that the self-reported anxiety measures were not collected as a time-series measure. In future studies, inspired by Xiao Yang et. al study [45], participants could be asked to rewatch video recordings of their Stressor 1 and 2 segments (each about 4 minutes long) and use a joystick to rate their anxiety level in a continuous manner. A casual relationship between the two contentious measures of breathing and the self-reported measure of anxiety during Stressor 2 could then be computed using Convergent Cross Mapping technique [44]. Should a casual relationship be found, the anxiety measure could be collapsed to a single data point and test for the mediation effect.

9.4 What is the relationship between the type of stressor and the skill of the user in breathing with the pacer?

Independent of the issue of the type of stressor, is practice using the pacer with the stressor important? In general, practice using a new technology is often important in being able to effectively use that technology. How much is needed? How persistent is the effect of practice—for example, would a third stressor block have shown a reduction in skin conductance by the treatment group? Sufficient practice may allow participants to master the act of matching their breathing with the pacer without using up cognitive resources required by the stressor task.

These questions suggest a $2 \times 3 \times 2$ table with stressor type (low versus high cognitive load), situation (in the moment versus, anticipatory, or after) as one axis and the experience level (novice versus expert) as the other axis. This study addresses the combination of novice users of the pacer experiencing an in-the moment, cognitively demanding stressor. In the future, it would be useful to sample from other combinations as well.

10 CONCLUSION AND FUTURE WORK

Our findings indicate that the breathing pacer led to expected change in breathing behavior. However, despite the documented efficacy of this pacer in decreasing self reported anxiety [17], and its success in modulating breathing, it did not produce reliable changes in skin conductance. We also failed to show that changes in self-reported anxiety were mediated by changes in breathing parameters.

We expected to find the pacer led to reductions in skin conductance, especially since the link between slow-paced breathing and reduced skin conductance is well established. Not finding the expected effect was surprising: we separately showed that the pacer modulates breathing behavior

and the breathing behavior unidirectionally, on a causal account, was linked to skin conductance. The experimental method was well designed and similar to previous studies [10, 12, 15] that used affective measures of skin conductance and self-reported anxiety to determine efficacy of a technological intervention. In addition, the experiment was adequately powered [46] to detect a medium size effect [9]. Instead, we observed that no matter how effective the pacer was in the absence of a stressor, or whether its usage began prior to engagement with stressor tasks, or whether we focused only on the active group, the psychophysiological calming effect of the pacer was indistinguishable from any other forms of affect regulation used by the control group.

These findings point to a puzzling situation: one measure of affect (skin conductance) suggested no differences between the group who fully modulated their breathing with the pacer, the group who didn't modulate their breathing with the pacer, and the group who neither received nor modulated their breathing with the pacer. The other affective measure (self-reported anxiety), on the other hand, showed differences between those who experienced the pacer (regardless of whether they modulated their breathing or not) and those who didn't. See Figure 7. Aggregating these two observations strongly suggests a placebo effect of the pacer for both active and passive groups.

However, such a conclusion strongly contradicts the following two findings: the pacer did modulate breathing and the breathing behavior, using a time-series analysis, was unidirectionally linked to skin conductance. The qualitative data on use of affect regulation during various stages of the study shed light into this paradoxical situation to suggest that presence of the spontaneous regulation in the control group and cognitive load to comply with the pacer in the treatment group could have contributed to not observing the effects of the interest.

Confronting our perplexing results, we have identified a set of considerations for future research in evaluating technology that assists in affect regulation. One consideration is to include the stressor anticipation phase in addition to the stressor phases. The difference in cognitive load between these two allow us to detect any masking effect the stressor could have on the affective measure of interest. Second, to include a collective of multiple measures of affect, in case one measure is masked by the cognitive load of the stressor choice, others could facilitate understanding the regulatory mechanism of the intervention. Another consideration is to collect qualitative data on the type of deployed affect regulation and collect measures that can verify them. Lastly, one should aim to understand the causal links between the technological interventions and affect in the context of the interaction among stressor types (e.g., low versus high cognitive load), situation (in the moment versus, anticipatory, or after), and user expertise (novice versus expert). Understanding how other technologies perform in this interaction space is important as well. More generally, further systematic research is needed to better understand

---

9. The minimal sample size was 39 to detect a medium effect size, with 80 percent power, as large as $d = 0.64$ which was observed in [12]. In our study both treatment and control groups had a sample size larger than 39.

the conditions under which a technological intervention will have impacts on an affective measure.

ACKNOWLEDGMENTS

REFERENCES

[1] J. J. Gross, "Emotion regulation: Affective, cognitive, and social consequences," *Psychophysiology*, vol. 39, no. 3, pp. 281–291, 2002.

[2] P. R. Goldin, K. McRae, W. Ramel, and J. J. Gross, "The neural bases of emotion regulation: reappraisal and suppression of negative emotion," *Biological psychiatry*, vol. 63, no. 6, pp. 577–586, 2008.

[3] J. J. Gross, "The emerging field of emotion regulation: an integrative review." *Review of general psychology*, vol. 2, no. 3, p. 271, 1998.

[4] J. J. Gross, H. Uusberg, and A. Uusberg, "Mental illness and well-being: an affect regulation perspective," *World Psychiatry*, vol. 18, no. 2, pp. 130–139, 2019.

[5] *Spire Stone*, 2018, https://spire.io/pages/stone. Accessed 8/31/2019.

[6] breathe App, https://support.apple.com/en-us/HT206999. Accessed 6/7/2019.

[7] *Lief Therapeutics*, 2018, https://getlief.com. Accessed 8/31/2019.

[8] *Vitali Sports Bra*, 2018, https://vitaliwear.com. Accessed 8/31/2019.

[9] *Lief Patch*, 2018, https://www.youtube.com/watch?v=aclzen-GGTE. Accessed 8/31/2019.

[10] P. E. Paredes, Y. Zhou, N. A.-H. Hamdan, S. Balters, E. Murnane, W. Ju, and J. A. Landay, "Just breathe: In-car interventions for guided slow breathing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, p. 28, 2018.

[11] J. Costa, A. T. Adams, M. F. Jung, F. Guimbetiere, and T. Choudhury, "Emotioncheck: leveraging bodily signals and false feedback to regulate our emotions," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 758–769.

[12] R. T. Azevedo, N. Bennett, A. Bilicki, J. Hooper, F. Markopoulou, and M. Tsakiris, "The calming effect of a new wearable device during the anticipation of public speech," *Scientific Reports*, vol. 7, 2017.

[13] A. Ghandeharioun and R. Picard, "Brightbeat: effortlessly influencing breathing for cultivating calmness and focus," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2017, pp. 1624–1631.

[14] E. N. Smith, E. Santoro, N. Moraveji, M. Susi, and A. J. Crum, "Integrating wearables in stress management

interventions: Promising evidence from a randomized trial." *International Journal of Stress Management*, 2019.

[15] J. Costa, F. Guimbretière, M. F. Jung, and T. Choudhury, "Boostmeup: Improving cognitive performance in the moment by unobtrusively regulating emotions with a smartwatch," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 2, p. 40, 2019.

[16] *Touchpoint*, https://thetouchpointsolution.com. Accessed 8/31/2019.

[17] P. Miri, E. Jusuf, A. Uusberg, H. Margarit, R. Flory, K. Isbister, K. Marzullo, and J. J. Gross, "Evaluating a personalizable, inconspicuous vibrotactile (piv) breathing pacer for in-the-moment affect regulation," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–12.

[18] K. Y. Choi and H. Ishii, "ambienbeat: Wrist-worn mobile tactile biofeedback for heart rate rhythmic regulation," in *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*, 2020, pp. 17–30.

[19] W. Chen, S. B. Oetomo, D. Tetteroo, F. Versteegh, T. Mamagkaki, M. S. Pereira, L. Janssen, and A. van Meurs, "Mimo pillow—an intelligent cushion designed with maternal heart beat vibrations for comforting newborn infants," *IEEE journal of biomedical and health informatics*, vol. 19, no. 3, pp. 979–985, 2014.

[20] P. M. Lehrer, E. Vaschillo, and B. Vaschillo, "Resonant frequency biofeedback training to increase cardiac variability: Rationale and manual for training," *Applied psychophysiology and biofeedback*, vol. 25, no. 3, pp. 177–191, 2000.

[21] E. G. Vaschillo, B. Vaschillo, and P. M. Lehrer, "Characteristics of resonance in heart rate variability stimulated by biofeedback," *Applied psychophysiology and biofeedback*, vol. 31, no. 2, pp. 129–142, 2006.

[22] G. G. Berntson, J. Thomas Bigger Jr, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone *et al.*, "Heart rate variability: origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, no. 6, pp. 623–648, 1997.

[23] G. Strauss-Blasche, M. Moser, M. Voica, D. McLeod, N. Klammer, and W. Marktl, "Relative timing of inspiration and expiration affects respiratory sinus arrhythmia," 2000.

[24] C. N. Joseph, C. Porta, G. Casucci, N. Casiraghi, M. Maffeis, M. Rossi, and L. Bernardi, "Slow breathing improves arterial baroreflex sensitivity and decreases blood pressure in essential hypertension," *hypertension*, vol. 46, no. 4, pp. 714–718, 2005.

[25] L. Bernardi, C. Porta, L. Spicuzza, J. Bellwon, G. Spadacini, A. W. Frey, L. Y. Yeung, J. E. Sanderson, R. Pedretti, and R. Tramarin, "Slow breathing increases arterial baroreflex sensitivity in patients with chronic heart failure," *Circulation*, vol. 105, no. 2, pp. 143–145, 2002.

[26] W. Kannel, H. Hubert, and E. Lew, "Vital capacity as a predictor of cardiovascular disease: the framingham study," *American heart journal*, vol. 105, no. 2, pp. 311–315, 1983.

[27] A. Zaccaro, A. Piarulli, M. Laurino, E. Garbella, D. Menicucci, B. Neri, and A. Gemignani, "How breath-control can change your life: a systematic review on psycho-physiological correlates of slow breathing. front hum neurosci. 2018; 12: 353," 2018.

[28] P. Lehrer, B. Vaschillo, T. Zucker, J. Graves, M. Katsamanis, M. Aviles, and F. Wamboldt, "Protocol for heart rate variability biofeedback training," *Biofeedback*, vol. 41, no. 3, pp. 98–109, 2013.

[29] I. Z. Khazan, *The clinical handbook of biofeedback: A step-by-step guide for training and practice with mindfulness.* John Wiley & Sons, 2013, iSBN: 978-1-118-48530-9.

[30] P. Miri, R. Flory, A. Uusberg, H. Culbertson, R. Harvey, A. Keman, E. Peper, and K. Gross, J. James Marzullo, "Piv: Placement, pattern, and personalization of an inconspicuous vibrotactile breathing pacer," 2019.

[31] S. Balters, E. L. Murnane, J. A. Landay, and P. E. Paredes, "Breath booster! exploring in-car, fast-paced breathing interventions to enhance driver arousal state," in *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2018, pp. 128–137.

[32] R. Goel, M. An, H. Alayrangues, A. Koneshloo, E. T. Lincoln, and P. E. Paredes, "Stress tracker—detecting acute stress from a trackpad: Controlled study," *Journal of Medical Internet Research*, vol. 22, no. 10, p. e22743, 2020.

[33] J. Van Dixhoorn, "Whole body breathing," *Principles and practices of stress management*, vol. 36, no. 3, pp. 291–332, 2007.

[34] *Lift,* http://tactilelabs.com/products/haptics/mmxc/. Accessed 6/7/2019.

[35] E. M. Bowden and M. Jung-Beeman, "Normative data for 144 compound remote associate problems," *Behavior Research Methods, Instruments, & Computers*, vol. 35, no. 4, pp. 634–639, 2003.

[36] M. J. Tobin, S. M. Guenther, W. Perez, R. F. Lodato, M. J. Mador, S. J. Allen, and D. R. Dantzker, "Konno-mead analysis of ribcage-abdominal motion during successful and unsuccessful trials of weaning from mechanical ventilation," *American Review of Respiratory Disease*, vol. 135, no. 6, pp. 1320–1328, 1987.

[37] K. Konno and J. Mead, "Measurement of the separate volume changes of rib cage and abdomen during breathing," *Journal of applied physiology*, vol. 22, no. 3, pp. 407–422, 1967.

[38] T. M. Marteau and H. Bekker, "The development of a six-item short-form of the state scale of the spielberger state—trait anxiety inventory (stai)," *British Journal of Clinical Psychology*, vol. 31, no. 3, pp. 301–306, 1992.

[39] S. Liu, R. X. Gao, J. Dinesh, J. Staudenmayer, and P. Freedson, "Tissue artifact removal from respiratory signals based on empirical mode decomposition," *Annals of Biomedical Engineering*, vol. 41, no. 5, pp. 1003–1015, 2013.

[40] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[41] A. Karagiannis, L. Loizou, and P. Constantinou, "Experimental respiratory signal analysis based on empirical mode decomposition," in *2008 First International Symposium on Applied Sciences on Biomedical and Communication Technologies*. IEEE, 2008, pp. 1–5.

[42] K. Gan, E. Yahyavi, and M. Ismail, "Contactless respiration rate measurement using optical method and empirical mode decomposition," *Technology and Health Care*, vol. 24, no. 5, pp. 761–768, 2016.

[43] Benedek and C. Kaernbac, "A continuous measure of phasic electrodermal activity," *Journal of Neuroscience Methods*, vol. 190, no. 1, pp. 80–91, 2010.

[44] G. Sugihara, R. May, H. Ye, C.-h. Hsieh, E. Deyle, M. Fogarty, and S. Munch, "Detecting causality in complex ecosystems," *science*, vol. 338, no. 6106, pp. 496–500, 2012.

[45] X. Yang, N. Ram, J. P. Lougheed, P. Molenaar, and T. Hollenstein, "Adolescents' emotion system dynamics: Network-based analysis of physiological and emotional experience." *Developmental psychology*, vol. 55, no. 9, p. 1982, 2019.

[46] M. Pardis, G. James J., M. Keith, U. Andero, and J. Emily, "*Pre-registered Hypothesis*," 2019, https://osf.io/epn3h.

Pardis Miri Pardis Miri is a Postdoctoral Scholar at Stanford University where she directs the WEHAB Lab.

Horia Margarit Horia Margarit is a Data Scientist and a Fellow at Spike Ventures.

Andero Uusberg Andero Uusberg is an Associate Professor of Affective Psychology at University of Tartu.

Keith Marzullo Keith Marzullo is a Professor and the Dean at the College of Information Studies at the University of Maryland.

Tali M. Ball Tali Ball is a clinical psychologist and senior research manager at Big Health.

Daniel Yamins Daniel Yamins is an Assistant Professor of Psychology and Computer Science at Stanford University.

Robert Flory Robert Flory is a hardware engineer at Intel Labs in Portland, OR.

James J. Gross James Gross is the Ernest R. Hilgard Professor of Psychology at Stanford University where he directs the Stanford Psychophysiology Laboratory.

APPENDIX A
BREATHING RATE MEASURE USING FFT

A simple approach to measure the breathing rate in a respiration signal is to identify and count all peaks in the signal and divide the count by the duration of the signal. This approach can be flawed when the respiration signal is not free from artifacts that may be incorrectly identified as breath peaks, but is often used because other methods present difficulties. In particular, while using a FFT is promising for determining the breathing rate (which should be the dominant frequency), outliers in a respiration signal, commonly formed due to posture change or change from active to passive breathing can result in the FFT computing a value close to zero as the dominant frequency, which is obviously an invalid breathing rate.

To more accurately identify dominant frequencies of a respiration signal using an FFT-based approach, we first removed artifacts as much as possible, using the generic EMD approach described in Section 6.1.1, and then identified outlier peaks and troughs as those which were more than three standard deviations away from the mean value of the peaks or valleys.[1] We then split the data into the segments between the outliers, and then performed an FFT on each segment to retrieve the dominant frequency for that segment. If no outlier was detected, we ran FFT on the entire outlier-free segment. We then computed the average of the frequencies over these segments to measure the breathing rate for the entire signal. However, this approach was still not fully robust, as there were several segments in which the FFT returned invalid values as the dominant frequency. These invalid values indicated that the cutoff we chose of 3 standard deviations for outlier identification would need to be further restricted. For those segments where this approach resulted in a valid measure, the value it produced was close to the value computed by the peak finding approach described in Section 6.1.2. Hence rather than continuing to adjust the outlier detection cutoff, we abandoned the FFT based approach in favor of the peak finding approach.

APPENDIX B
MODEL-BASED APPROACHES TO SKIN CONDUCTANCE ANALYSIS

The mapping between observed skin conductance (SC) and sympathetic arousal (SA) is known as skin conductance analysis. This analysis uses two general approaches: a model-based approach which uses mathematical models to map between observed SC and SA, and a model-free approach, such as the average mean of an observed SC within a certain time frame. The model-free approach is simple but is limited because it does not account for overlapping skin conductance responses (SCRs). A model-based approach addresses this limitation by first mapping SC data to sudomotor nerve (SN) activity, which contains discrete peaks that do not overlap, and then, inferring SA from SN activity.

Despite the complexity of model-based approaches, they are generally preferred to model-free approaches because they provide more accurate estimations of stress responses.

Among model-based approaches, Ledalab [1] is one of the widely used packages to infer SA from observed SC. It infers SA from observed SC in two steps: inference of SN from SC, followed by inference of SA levels from SN activity. This relation can be expressed as $SA \rightarrow SN \rightarrow SC$. Ledalab uses a casual operation to estimate sudomotor nerve activity from skin conductance and then uses a peak detection method to infer sympathetic arousal from the estimated sudomotor nerve activity. The casual operation inferring sudomotor nerve activity, involves two consecutive steps. The first step is a deconvolution of the observed skin conductance with an impulse response function, which is an estimate of a single skin conductance response[2]. The deconvolution results in a *driver* signal that contains both tonic (slow-changing) and phasic (fast-changing) components. The second step is an estimation of the tonic activity from the driver in four consecutive sub-steps: (1) smoothing the driver signal using a convolution with a Gaussian function; (2) identifying peaks and local minimums before and after each pick in the smoothed driver (any points that are not within a peak section are considered as part of the tonic driver); (3) interpolating missing data values to create a continuous tonic signal; and (4) subtracting the tonic signal from the original driver to calculate the phasic component of the driver. The derived phasic component ("phasic driver") is used as an estimation of sudomotor nerve activity [1].

To infer sympathetic arousal, Ledalab detects peaks in the phasic driver above the 0.01 microsiemens (or any other threshold the user may wish to set based on the desired sensitivity). Ledalab then presents three estimations for sympathetic arousal: CDA.AmpSum which reports the sum of the amplitudes of the peaks in the phasic driver during a response window, CDA.nSCR which represents the total number of significant (above the threshold) peaks in a response window, and CDA.SCR which represents the average phasic driver value in a response window. [2] The first and third measures are in microsiemens, and the second is a cardinal number.

APPENDIX C
OTHER MEASURES OF BREATHING MODULATION ANALYSIS

After failing to show the mediation effect of breathing rate and active breathing ratio in explaining the link between the pacer and skin conductance (during the last one minute), we further tested whether the other modulated measures of breathing could explain the pacer's effect as a mediator. We found that the following two measures of breathing were modulated by the pacer intervention: Breathing irregularity and chest to abdomen displacement ratio. The degree to which the participants followed the pacer varied, which resulted in more inconsistent breathing (measured through coefficient of variance) and decreased ratio of abdomen to chest breathing (slope of a regression line fitted over a

---

1. To extract the peaks related information, we used Python `find_peaks` with the following settings: `find_peaks(signal, distance = 2 × sampling_frequency, width = .8 × sampling_frequency, prominence = .3)`.

2. The impulse response function formula is $e^{(-t/2)} - e^{(-t/.75)}t = 0 : .01 : 20$
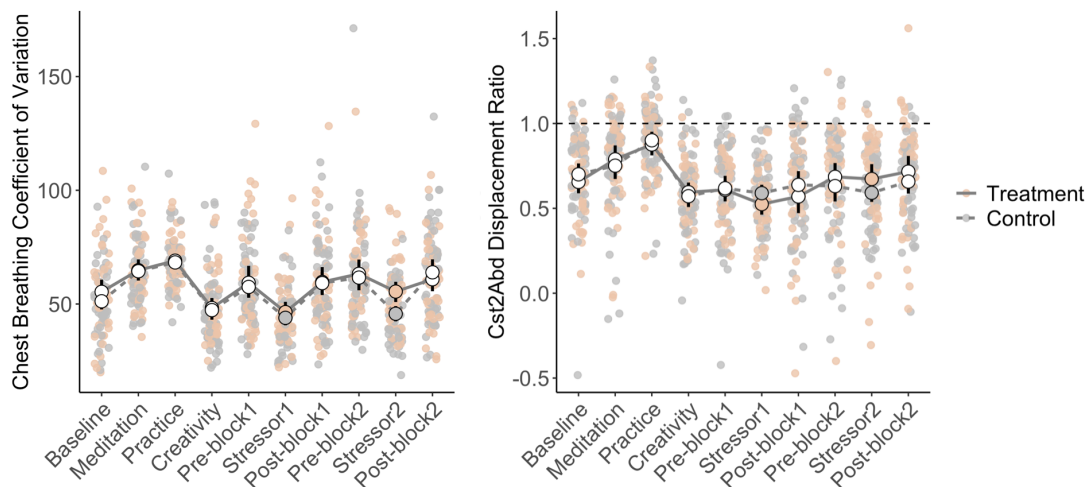
Fig. 1: **Left:** Average breathing irregularity (coefficient of variance) during the study for both treatment and control groups. An interaction effect was observed between group (treatment and control) and condition (Stressor 1 and Stressor 2): the treatment group receiving vibrotactile patterns during Stressor 2 experienced an increase in chest breathing irregularity compared to the control group (see colored dots). Solid line indicates treatment group; dotted line indicates control group. **Right:** Chest to abdomen displacement ratio during the study for both treatment and control groups. An interaction effect was observed between group (treatment and control) and condition (Stressor 1 and Stressor 2): the treatment group receiving vibrotactile patterns during Stressor 2 experienced an increase in chest breathing compared to the control group (see colored dots). Solid line indicates treatment group; dotted line indicates control group.

Konno-Mead [3, 4] plot). These results are explained in more details below.

C.1 Efficacy of the pacer in reducing breathing irregularity

To measure the irregularity in breathing depth, we computed the coefficient of variance (CV) of the respiration signals, which is the standard deviation expressed as a percentage of a signal mean (i.e., $\frac{SD}{Mean} \times 100$). A higher CV corresponds to more irregularity in the depth of breathing.

We hypothesized an interaction effect between group (treatment and control) and condition (Stressor 1 and Stressor 2) for the dependent variable of breathing irregularity. Specifically, we expected to observe less breathing irregularities from Stressor 1 to Stressor 2 in the treatment group, but no change in the control group. We thought that the pacer would contribute to active breathing (more regular breathing) while participants were performing stressor tasks.

We observed an interaction effect between group (treatment and control) and condition (Stressor 1 and Stressor 2) for breathing irregularity (intercept = 46.43, $\beta = -7.25$, p = .002, CI = $[-11.64, -2.74]$). However, in contrast to our hypothesis, this was characterized by an increase rather than a decrease in breathing irregularity from Stressor 1 to Stressor 2 in the treatment group (shown in Figure 1, left). Although this is inconsistent with our original hypothesis, it supports our earlier observation that participants switched between passive and active breathing rather than consistently performing active slow-paced breathing while the pacer was active. Such switching behavior would, of course, result in a larger variance in the breathing signals.

C.2 Efficacy of the pacer in reducing chest to abdomen displacement ratio

To measure the ratio of chest-to-abdomen displacement, we examined ribcage-abdominal motion using the method of Konno and Mead [3, 4], where the extensions of the abdomen and rib cage are displayed on an XY plot. There is no axis corresponding to time on a Konno-Mead plot; instead, the breathing waves together generate a trace on the Cartesian planeXY plot. When the chest and abdomen move relatively in synchronization, the trace is a figure that slopes upwards and to the right. We then fit in an ordinary least squares (OLS) regression line which is a measure of the ratio of abdomen to chest motion: slopes less than 1 result from more abdominal compared to chest breathing. Figure 2 shows Konno-Mead plots for participant subject `s021` during Stressor 1 and Stressor 2. This participant was in the treatment group and so received the pacer intervention during Stressor 2. The figure shows that the participant exhibited more abdominal breathing during Stressor 2 as compared to Stressor 1.

We hypothesized an interaction effect between group (treatment and control) and condition (Stressor 1 and Stressor 2) for the dependent variable of chest-to-abdomen displacement ratio. Specifically, we expected to observe no change in the chest-to-abdomen displacement ratio in the control group, but more abdominal breathing in the treatment group during Stressor 2 when the pacer was active. We hypothesized more abdominal breathing when the pacer was active for two reasons: early in the protocol, [3] participants practiced abdominal breathing; and the tactors deployed patterns of vibrations on the abdomen area which

---

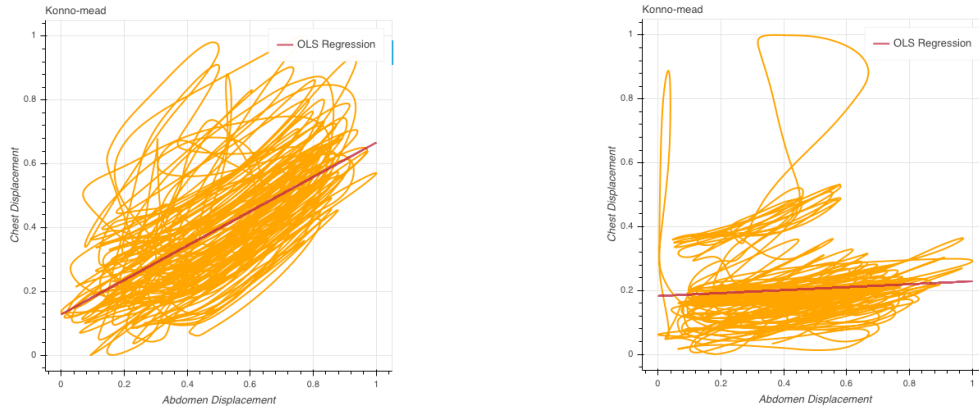3. See details of the experimental design in the protocol section.

Fig. 2: Examples of Konno-Mead plots for two participants in the treatment group while the pacer was active during the Stressor 2 condition. The participant on the left engaged in more chest breathing as compared to the participant on the right, indicated with the red regression line.

could engage the tendency of more abdominal slow-paced breathing [5, 6].

We had expected that when participants switched from passive to active breathing, they would demonstrate more abdominal displacement. This is because the participants practiced abdominal breathing in an earlier stage of the study. Inconsistent with our expectation, when participants switched between passive and active breathing, they demonstrated more chest than abdominal displacement (intercept = −.75, $\beta = -.34$, $p = .03$, CI = $[-.64, -.03]$). The observed effect was in the opposite direction of what we had anticipated (See Figure 1, right).

APPENDIX D

TESTING THE CAUSAL LINK BETWEEN BREATHING AND SKIN CONDUCTANCE USING CONVERGENT CROSS MAPPING

We tested the causal relationship between breathing and skin conductance for each participant using Convergent Cross Mapping (CCM) [7], which is a method that distinguishes causality from correlation in nonseparable weakly connected dynamic systems. The more commonly used approach in time-series analysis is based on Granger causality [8], which is best suited for purely systems where the influences of the causal variables are separable (independent of each other). CCM, on the other hand, is based on the theory of dynamic systems where causal variables have synergistic effects. Granger causality states that the predictability of Y declines when X was removed. That is if X is removed, its information will be eliminated at the same time, which suggests separability – the information for a causative factor only depends on one variable. However, the assumption of separability is appropriate to the stochastic systems or linear systems because separability implies that the system is decomposable into separable entities with no dependencies between them.

CCM does not require this kind of separability. CCM is based on Takens' embedding theorem [9], which states that by means of a lagged reconstruction of a single variable, we can obtain a structure (i.e., shadow manifold) that is topologically equivalent to the manifold attractor of the

dynamic system of interest. That is, the shadow manifold is diffeomorphic — it has a one-to-one mapping to the main manifold. Therefore, if two variables belong to the same dynamic system, each of their shadow manifolds will also be diffeomorphic. Time points that are nearby on the one of the shadow manifolds will correspond to points that are nearby on the other shadow manifold. Therefore, we can use nearby points on one shadow manifold to identify whether correspondent points on the other shadow manifold are close by. If so, then the variables are causally related. For example, if breathing signal forces skin conductance unidirectionally, skin conductance will contain information about breathing but not vice versa. Consequently, the shadow manifold of breathing can be predicted from the shadow manifold of skin conductance, but not the other way around. This technique is called cross mapping. With longer time series, the shadow manifolds are denser and so the nearest neighbors are closer which leads to increased precision of the cross mapping estimates. This phenomenon is called convergent cross mapping. We have applied CCM to the two time-series of breathing (chest or abdominal displacement) and corresponding skin conductance, and have shown that, indeed, breathing is causing the skin conductance unidirectionally. We observed this causal relationship for each participant regardless of group membership (Treatment or Control) or block (Stressor or Stressor2).

APPENDIX E

VALIDITY OF SKIN CONDUCTANCE MEASURE

First, we investigated the validity of the skin conductance data. Figure 8 shows that the skin conductance values were within the expected ranges and they dropped, as expected, during the V-breathing Practice condition. This suggests that the skin conductance data obtained in this study were valid and reliable as they revealed a well-known link between slow paced breathing and drop in skin conductance.

Second, we investigated whether participants experienced sympathetic arousal as an indication of stress during Stressor 1 and Stressor 2. Consistent with our expectation, the large number of peaks in the phasic drive (`CDA.nSCR`) during Stressor 1 and Stressor 2 demonstrated that the

participants experienced significantly higher levels of sympathetic arousal during the stressor tasks segments as compared to the rest of the study conditions. This suggests that the CRA tasks were indeed stressful for the participants (Figure 8, right).

Third, we investigated whether there was a habituation effect to the stressor when participants began Stressor 2. Not surprisingly, we observed a habituation effect from Stressor 1 to Stressor 2 in both the treatment and control groups. A main effect of condition on all three measures of `CDA.SCR` (intercept = .39 , $\beta = -.08$, $p < .001$, CI = $[-.08, .0007]$), `CDA.nSCR` (intercept = 54.86, $\beta = -13.90$, $p < .001$, CI = $[-13.90, -.04]$), and `CS.mean` (intercept = 11.55, $\beta = -.82$, $p < .001$, CI = $[-1.15, -.48]$) were observed, which indicates that the average observed skin conductance and the inferred sympathetic arousal both decreased from the Stressor 1 to Stressor 2 condition, regardless of the group.

Taken together, these findings show that the changes in skin conductance measure and the experimental design were meaningful.

APPENDIX F
OTHER EXPLANATIONS FOR NOT OBSERVING THE LINK BETWEEN THE SKIN CONDUCTANCE AND THE PACER

The following investigations into a link between skin conductance and using the pacer are more speculative than those described in Section 8.3.

It is possible that focusing on the entire 4-minute window size for the stressors may have contributed to the lack of the anticipated interaction effect. Therefore, we focused on the first minute or the last minute window of the stressors for the following reasons. During Pre-stressor 2, the participants practiced slow-paced breathing, which could give the treatment group a better head start in keeping their breathing rate low, which in turn would lead to lower skin conductance during the early interaction with the stressor. And, during the last minute of the stressor, the treatment group participant has had some time to acquaint themselves with following the pacer while performing CRA tasks.

We investigated this further by testing for an interaction effect at the beginning and at the end of the skin conductance window for the measure of CS.mean. With a window of the first minute of the stressors, we failed to observe an interaction effect. Therefore, the early start of slow-paced breathing, prior engagement with the stressor tasks, did not enhance the skin conductance drop during the first minute window of the stressors. Note that the average level of skin conductance (CS.mean) for both treatment and control groups were indistinguishable during the two minutes of Pre-stressor 2 (see Figure 8).

With a window of the final minute of the stressor, we observed an interaction effect (intercept = 11.12, β = -.60, p = .01, CI = $[-1.04, -.16,]$) between group (treatment and control) and condition (Stressor 1 and Stressor 2). This suggests that participants benefited from becoming acquainted with using the pacer while faced with time-pressured CRA tasks. Perhaps the first three minutes can be thought of as a practice period before the final minute's performance, and the effect of active breathing on reducing skin conductance

became more visible during the final minute. With such practice, the CRA task-induced component that increases cognitive load and consequently skin conductance was weakened compared to the stress reduction component of active breathing that reduces skin conductance.

Because we found an interaction effect suggesting that skin conductance was reduced in the treatment group more than in the control group (only in the last minute of of four-minute window), we ran the mediation analysis for the window of the last minute, and found the pacer effect in reducing skin conductance calculated through global mean (see Section 8.3). However, we then failed to find the mediation effect that could explain a casual path between vibrotactile patterns of the pacer and skin conductance via breathing rate. We did observe a direct effect and a total effect that links the two together (ACME =.03, CI = $[-.21, -.29]$, $p = .76$; ADE = $-.59$, CI = $[-1.16, -.02]$, $p = .04$; total effect = $-.55$, CI = $[-1.02, -.11]$, $p = .01$). This means that there was a drop in skin conductance from Stressor 1 to Stressor 2, most likely due to habituation effect, but it was not explainable through breathing rate.

In sum, despite separately observing pacer causal impact on breathing, and breathing casual impact on skin conductance, we couldn't demonstrate this three-step relationship using a mediation analysis.

Another plausible explanation is that this study was underpowered. We did our original power analysis using the effect size (Cohen's d) in Dopple [10], which we computed to be 0.64 for the measure of SC.mean. However, in this study we observed the effect size of Cohen's $d = .3$ in the difference in CDA.nSCR between active and control group participants (t = -1, df = 21, p = 0.33). To reliably detect a difference as small as 0.3, 176 participants in each group would be needed. Given the current sample size of 17 in active and 52 in treatment groups, there is about a 13% probability of detecting true differences as small as 0.3. [4] We focused on comparison between the two groups of active and control instead of treatment and control because the active group were the ideal representation of the treatment group in which everyone made an attempt to follow the pacer and succeed. Therefore, we expect that they would benefit from slow paced breathing and produce a pure effect size. [5]

This explanation is unlikely to be the only reason because, in the absence of a cognitive stressor, the effect of the pacer was very large. The block effect (Cohen's d) from condition Baseline to Slow-paced breathing for both active and control groups, for CDA.nSCR was 1.22 and significant (t = 7.19, df = 80, $p < .0001$). With the introduction of the stress, the effect size dropped down to to 0.3 as described above. Thus, the presence of the cognitive stressor was probably not a contributing factor regardless of the study being underpowered. Supporting this argument, the study

---

4. We used R package "effectsize" to convert the F test result in Dopple [10] study to Cohen's d: effect_size = F_to_d(5.23, 1, 50, paired = F);pwr.t.test(d = effect_size, power = 0.8, sig.level = 0.05, type = "two.sample")

5. We reported the effect sizes for CDA.nSCR instead of SC.mean that was reported in the Dopple paper, because between the three measures of skin conductance (CDA.SCR, CDA.nSCR, and CS.mean), CDA.nSCR provided the largest effect size.The effect size of CS.mean was negligible (ranged between .03 to .09) compared to CDA.nSCR.

reported in [11] also found that the effect size of slow paced breathing was around .2 and was insignificant, in the presence of a cognitive stressor.

It is plausible that the stressor task had a masking effect. The creativity tasks used in this study were cognitively intensive and required a significant amount of attention. [6] High cognitive load has been associated with an increase in skin conductance response [12]. Therefore, the cognitive demands of the CRA tasks may have induced a strong upward force on skin conductance that overshadowed the stress reduction effect we were hoping to observe. Indeed, we found that the performance in the CRA did not differ between group (treatment and control) and condition (Stressor 1 and Stressor 2) suggesting that the cognitive contribution to the skin conductance signal remained constant. It is possible that this effect on skin conductance was strong enough that any reduction to stress was not detectable by the skin conductance signal. This explanation is unlikely because we observed that skin conductance was indistinguishable between the treatment and control group, during Pre-stressor 2.

APPENDIX G
ALGORITHM FOR TESTING THE MEDIATION EFFECT OF BREATHING RATE CHANGE TO EXPLAIN DROP IN SKIN CONDUCTANCE

To test the confirmatory analyses of the links between the pacer's patterns of vibrations and the independent variables of self-reported anxiety and skin conductance, mediated through breathing rate, we used the mediate function of the mediation package in R. Algorithm 1 gives the pseudocode of the algorithm we used.

Algorithm 1: Testing the mediation effect of breathing rate change to explain drop in skin conductance

```
# Breathing Rate Change model
bpm.model = lm(BPM2 ~ BPM1, data = df)
# Exclude the NAs
bpm.model.no.NA = update(bpm.model, na.action =
  na.exclude)
# Add column of residuals to the data
df["BPM_change"] = resid(bpm.model.no.NA)
#————————————————
# Skin Conductance Change model
sc.model = lm(SC2 ~ SC1, data = df)
# Exclude the NAs
sc.model.no.NA = update(sc.model, na.action =
  na.exclude)
# Add column of residuals to the data
df["SC_change"] = resid(sc.model.no.NA)
#————————————————
# Mediator model
M.model = lm(BPM_change ~ group , data = df)
#————————————————
# Predictor model
Y.model = lm(SC_change ~ group + BPM_change,
  data = df)
#————————————————
# Bootstrap the mediation effect
fitMedBoot = mediation::mediate(M.model ,Y.model
  , boot=TRUE, sims=1000, treat="group",
  mediator="BPM_change", control.value =
  "Control", treat.value = "Treatment")
summary(fitMedBoot)
plot(fitMedBoot)
```

REFERENCES

[1] Benedek and C. Kaernbac, "A continuous measure of phasic electrodermal activity," *Journal of Neuroscience Methods*, vol. 190, no. 1, pp. 80–91, 2010.

[2] *LedaLab*, 2018, http://www.ledalab.de/. Accessed 8/31/2020.

[3] M. J. Tobin, S. M. Guenther, W. Perez, R. F. Lodato, M. J. Mador, S. J. Allen, and D. R. Dantzker, "Konno-mead analysis of ribcage-abdominal motion during successful and unsuccessful trials of weaning from mechanical ventilation," *American Review of Respiratory Disease*, vol. 135, no. 6, pp. 1320–1328, 1987.

[4] K. Konno and J. Mead, "Measurement of the separate volume changes of rib cage and abdomen during breathing," *Journal of applied physiology*, vol. 22, no. 3, pp. 407–422, 1967.

[5] L. Narici, I. Modena, R.-J. Opsomer, V. Pizzella, G. Romani, G. Torrioli, R. Traversa, and P. Rossini, "Neuromagnetic somatosensory homunculus: a non-invasive approach in humans," *Neuroscience letters*, vol. 121, no. 1-2, pp. 51–54, 1991.

[6] J. D. Brown, R. B. Gillespie, D. Gardner, and E. A. Gansallo, "Co-location of force and action improves identification of force-displacement features," in *Haptics Symposium (HAPTICS), 2012 IEEE*. IEEE, 2012, pp. 187–193.

[7] G. Sugihara, R. May, H. Ye, C.-h. Hsieh, E. Deyle, M. Fogarty, and S. Munch, "Detecting causality in complex ecosystems," *science*, vol. 338, no. 6106, pp. 496–500, 2012.

[8] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: journal of the Econometric Society*, pp. 424–438, 1969.

[9] L. Noakes, "The takens embedding theorem," *International Journal of Bifurcation and Chaos*, vol. 1, no. 04, pp. 867–872, 1991.

[10] R. T. Azevedo, N. Bennett, A. Bilicki, J. Hooper, F. Markopoulou, and M. Tsakiris, "The calming effect of a new wearable device during the anticipation of public speech," *Scientific Reports*, vol. 7, 2017.

[11] V. Rauschel, A. Straube, F. Süß, and R. Ruscheweyh, "Responsiveness of the autonomic nervous system during paced breathing and mental stress in migraine patients," *The journal of headache and pain*, vol. 16, no. 1, pp. 1–10, 2015.

[12] N. Nourbakhsh, F. Chen, Y. Wang, and R. A. Calvo,

---

6. In addition, the nine second timer, the loud announcement of "correct"' and "incorrect" as feedback at the end of each task, and the nature of the compound association of the tasks had all contributed to stress increase.

"Detecting users' cognitive load by galvanic skin response with affective interference," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 7, no. 3, pp. 1–20, 2017.